



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification: C12N 15/29, C12N 15/82, C12P 21/02, C12Q 1/68	A2	(11) International Publication Number: WO 00/28036 (43) International Publication Date: 18 May 2000 (18.05.2000)
(21) International Application Number: PCT/US99/25953 (22) International Filing Date: 04 November 1999 (04.11.1999) (30) Priority Data: 60/107,242 05 November 1998 (05.11.1998) US (60) Parent Application or Grant E.I. DU PONT DE NEMOURS AND COMPANY [/]; (). FAMODU, Omolayo, O. [/]; (). FANG, Yiwen [/]; (). LIU, Zhan-Bin [/]; (). MIAO, Guo-Hua [/]; (). ODELL, Joan, T. [/]; (). FAMODU, Omolayo, O. [/]; (). FANG, Yiwen [/]; (). LIU, Zhan-Bin [/]; (). MIAO, Guo-Hua [/]; (). ODELL, Joan, T. [/]; (). FEULNER, Gregory, J. ; ().		Published
(54) Title: DISEASE RESISTANCE FACTORS (54) Titre: FACTEURS DE RESISTANCE A LA MALADIE (57) Abstract <p>This invention relates to an isolated nucleic acid fragment encoding an NPR1 gene. The invention also relates to the construction of a chimeric gene encoding all or a portion of the NPR1, in sense or antisense orientation, wherein expression of the chimeric gene results in production of altered levels of the NPR1 in a transformed host cell.</p> (57) Abrégé <p>La présente invention concerne un fragment d'acide nucléique isolé codant un gène NPR1. L'invention concerne également la construction d'un gène chimérique codant tout ou partie du NPR1, en orientation sens ou antisens, auquel cas, l'expression du gène chimérique aboutit à une modification des niveaux de production du NPR1 dans une cellule hôte transformée.</p>		

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C12N 15/29, 15/82, C12Q 1/68, C12P 21/02	A2	(11) International Publication Number: WO 00/28036 (43) International Publication Date: 18 May 2000 (18.05.00)
(21) International Application Number: PCT/US99/25953 (22) International Filing Date: 4 November 1999 (04.11.99) (30) Priority Data: 60/107,242 5 November 1998 (05.11.98) US (71) Applicant (for all designated States except US): E.I. DU PONT DE NEMOURS AND COMPANY [US/US]; 1007 Market Street, Wilmington, DE 19898 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): FAMODU, Omolayo, O. [US/US]; 216 Barrett Run Place, Newark, DE 19702 (US). FANG, Yiwen [CN/US]; 203 Louis Lane, Hockessin, DE 19707 (US). LIU, Zhan-Bin [CN/US]; 108C Presidential Drive, Greenville, DE 19807 (US). MIAO, Guo-Hua [CN/US]; 202 Cherry Blossom Place, Hockessin, DE 19707 (US). ODELL, Joan, T. [US/US]; P.O. Box 826, Unionville, PA 19375 (US). (74) Agent: FEULNER, Gregory, J.; E.I. du Pont de Nemours and Company, Legal Patent Records Center, 1007 Market Street, Wilmington, DE 19898 (US).		(81) Designated States: AE, AL, AU, BA, BB, BG, BR, CA, CN, CR, CU, CZ, DM, EE, GD, GE, HR, HU, ID, IL, IN, IS, JP, KP, KR, LC, LK, LR, LT, LV, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, SL, TR, TT, UA, US, UZ, VN, YU, ZA, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>Without international search report and to be republished upon receipt of that report.</i>
(54) Title: DISEASE RESISTANCE FACTORS (57) Abstract This invention relates to an isolated nucleic acid fragment encoding an NPR1 gene. The invention also relates to the construction of a chimeric gene encoding all or a portion of the NPR1, in sense or antisense orientation, wherein expression of the chimeric gene results in production of altered levels of the NPR1 in a transformed host cell.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

Description

5

10

15

20

25

30

35

40

45

50

55

TITLE

DISEASE RESISTANCE FACTORS

This application claims the benefit of U.S. Provisional Application No. 60/107,242, filed November 5, 1998.

FIELD OF THE INVENTION

This invention is in the field of plant molecular biology. More specifically, this invention pertains to nucleic acid fragments encoding NPR1 in plants and seeds.

BACKGROUND OF THE INVENTION

Pathogens annually cause billions of dollars in damage to crops worldwide.

Consequently, an increasing amount of research has been dedicated to developing novel methods for controlling plant diseases. Such studies have centered on the plant's innate ability to resist pathogen invasion in an effort to support the plant's own defenses to counter pathogen attacks (Staskawicz et al. (1995) *Science* 268:661-667; Baker et al. (1997) *Science* 276:726-733). One such defense mechanism under study is known as systemic acquired resistance (SAR; reviewed in Ryals et al. (1996) *Plant Cell* 8:1809-1819). SAR is defined as a generalized defense response, which is often induced by avirulent pathogens and provides enhanced resistance to a broad spectrum of virulent pathogens. Avirulent pathogens carry an avirulence (avr) gene whose product can be recognized by the product of a corresponding resistance (R) gene carried by plants. Such recognition triggers both a programmed cell death response, known as the hypersensitive response (HR), around the point of pathogen infection and release of a systemic SAR-inducing signal (Hammond-Kosack and Jones (1996) *Plant Cell* 8:1773-1791). After a rapid and localized HR, the elevated state of resistance associated with SAR is effective throughout the plant for a period of time ranging from several days to a few weeks. Coinciding with the onset of SAR is the transcriptional activation of the pathogenesis-related (PR) genes. These genes encode proteins that exhibit antimicrobial activities (Ward et al. (1991) *Plant Cell* 3:1085-1094).

In *Arabidopsis*, expression of PR-1, β -1,3-glucanase (BGL2), and PR-5 has been shown to be tightly correlated with resistance to virulent bacterial, fungal, and oomycete pathogens; therefore, these genes are used as molecular markers for SAR (Uknes et al. (1992) *Plant Cell* 4:645-656). The *Arabidopsis* NPR1 controls the onset of SAR. Mutants with defects in NPR1 fail to respond to various SAR-inducing treatments, displaying little expression of PR genes and exhibiting increased susceptibility to infections. NPR1 was cloned using a map-based approach and was found to encode a novel protein containing ankyrin repeats. The lesion in one *npr1* mutant allele disrupted the ankyrin consensus sequence, suggesting that these repeats are important for NPR1 function. Furthermore, transformation of the cloned wild-type *npr1* gene into *npr1* mutants not only complemented the mutations, restoring the responsiveness to SAR induction with respect to PR-gene expression and resistance to infections, but also rendered the transgenic plants more resistant

to infection by *P. syringae* in the absence of SAR induction (Cao et al. (1997) *Cell* 88:57-63).

SUMMARY OF THE INVENTION

The present invention relates to isolated polynucleotides comprising a nucleotide sequence encoding a first polypeptide of at least 80 amino acids that has at least 75% identity based on the Clustal method of alignment when compared to a polypeptide selected from the group consisting of a corn NPR1 polypeptide of SEQ ID NO:2, a rice NPR1 polypeptide of SEQ ID NO:4 and SEQ ID NO:6, and a wheat NPR1 polypeptide of SEQ ID NO:8. The present invention also relates to an isolated polynucleotide comprising the complement of the nucleotide sequences described above.

It is preferred that the isolated polynucleotide of the claimed invention consists of a nucleic acid sequence selected from the group consisting of SEQ ID NOS:1, 3, 5, and 7 that codes for the polypeptide selected from the group consisting of SEQ ID NOS:2, 4, 6, and 8. The present invention also relates to an isolated polynucleotide comprising a nucleotide sequences of at least one of 40 (preferably at least one of 30) contiguous nucleotides derived from a nucleotide sequence selected from the group consisting of SEQ ID NOS:1, 3, 5, 7, and the complement of such nucleotide sequences.

The present invention relates to a chimeric gene comprising an isolated polynucleotide of the present invention operably linked to suitable regulatory sequences.

The present invention relates to an isolated host cell comprising a chimeric gene of the present invention or an isolated polynucleotide of the present invention. The host cell may be eukaryotic, such as a yeast or a plant cell, or prokaryotic, such as a bacterial cell. The present invention also relates to a virus, preferably a baculovirus, comprising an isolated polynucleotide of the present invention or a chimeric gene of the present invention.

The present invention relates to a process for producing an isolated host cell comprising a chimeric gene of the present invention or an isolated polynucleotide of the present invention, the process comprising either transforming or transfecting an isolated compatible host cell with a chimeric gene or isolated polynucleotide of the present invention.

The present invention relates to an NPR1 polypeptide of at least 80 amino acids comprising at least 75% identity based on the Clustal method of alignment compared to a polypeptide selected from the group consisting of SEQ ID NOS:2, 4, 6, and 8.

The present invention relates to a method of selecting an isolated polynucleotide that affects the level of expression of an NPR1 polypeptide in a plant cell, the method comprising the steps of:

- constructing an isolated polynucleotide of the present invention or an isolated chimeric gene of the present invention;
- introducing the isolated polynucleotide or the isolated chimeric gene into a host cell, preferably a plant cell;

5 measuring the level of an NPR1 polypeptide in the host cell containing the
isolated polynucleotide; and
10 comparing the level of an NPR1 polypeptide in the host cell containing the
isolated polynucleotide with the level of an NPR1 polypeptide in a host cell that does not
5 contain the isolated polynucleotide.

The present invention relates to a method of obtaining a nucleic acid fragment
encoding a substantial portion of an NPR1 polypeptide gene, preferably a plant NPR1
15 polypeptide gene, comprising the steps of: synthesizing an oligonucleotide primer
comprising a nucleotide sequence of at least one of 40 (preferably at least one of 30)
10 contiguous nucleotides derived from a nucleotide sequence selected from the group
consisting of SEQ ID NOs:1, 3, 5, 7, and the complement of such nucleotide sequences; and
20 amplifying a nucleic acid fragment (preferably a cDNA inserted in a cloning vector) using
the oligonucleotide primer. The amplified nucleic acid fragment preferably will encode a
portion of an NPR1 amino acid sequence.

15 The present invention also relates to a method of obtaining a nucleic acid fragment
encoding all or a substantial portion of the amino acid sequence encoding an NPR1
25 polypeptide comprising the steps of: probing a cDNA or genomic library with an isolated
polynucleotide of the present invention; identifying a DNA clone that hybridizes with an
isolated polynucleotide of the present invention; isolating the identified DNA clone; and
20 sequencing the cDNA or genomic fragment that comprises the isolated DNA clone.

30 The present invention also relates to an expression cassette comprising at least one
isolated polynucleotide of the present invention operably linked to a promoter.

The present invention also relates to a method for positive selection of a transformed
cell comprising:

- 35 25 (a) transforming a plant cell, preferably a monocot, with an expression cassette
of the present invention; and
(b) growing the transformed plant cell under conditions allowing expression of
40 the polynucleotide in an amount sufficient to induce disease resistance to provide a positive
selective means.

30 BRIEF DESCRIPTION OF THE SEQUENCE LISTING

The invention can be more fully understood from the following detailed description
and the accompanying Sequence Listing which form a part of this application.

45 Table 1 lists the polypeptides that are described herein, the designation of the cDNA
clones that comprise the nucleic acid fragments encoding polypeptides representing all or a
35 substantial portion of these polypeptides, and the corresponding identifier (SEQ ID NO:) as
used in the attached Sequence Listing. The sequence descriptions and Sequence Listing
50 attached hereto comply with the rules governing nucleotide and/or amino acid sequence
disclosures in patent applications as set forth in 37 C.F.R. §1.821-1.825.

TABLE 1

Disease Resistance Factors

Protein	Clone Designation	SEQ ID NO:	
		(Nucleotide)	(Amino Acid)
Corn NPR1	p0006.cbyvc82rx	1	2
Rice NPR1	rl0n.pk0063.d10	3	4
Rice NPR1	rr1.pk0001.a11	5	6
Wheat NPR1	wre1n.pk0122.c2	7	8

The Sequence Listing contains the one letter code for nucleotide sequence characters and the three letter codes for amino acids as defined in conformity with the IUPAC-IUBMB standards described in *Nucleic Acids Res.* 13:3021-3030 (1985) and in the *Biochemical J.* 219 (No. 2):345-373 (1984) which are herein incorporated by reference. The symbols and format used for nucleotide and amino acid sequence data comply with the rules set forth in 37 C.F.R. §1.822.

DETAILED DESCRIPTION OF THE INVENTION

In the context of this disclosure, a number of terms shall be utilized. As used herein, a "polynucleotide" is a nucleotide sequence such as a nucleic acid fragment. A polynucleotide may be a polymer of RNA or DNA that is single- or double-stranded, that optionally contains synthetic, non-natural or altered nucleotide bases. A polynucleotide in the form of a polymer of DNA may be comprised of one or more segments of cDNA, genomic DNA, or synthetic DNA. An isolated polynucleotide of the present invention may include at least one of 40 contiguous nucleotides, preferably at least one of 30 contiguous nucleotides, most preferably one of at least 15 contiguous nucleotides, of the nucleic acid sequence of the SEQ ID NOs:1, 3, 5, and 7.

As used herein, "substantially similar" refers to nucleic acid fragments wherein changes in one or more nucleotide bases results in substitution of one or more amino acids, but do not affect the functional properties of the polypeptide encoded by the nucleotide sequence. "Substantially similar" also refers to nucleic acid fragments wherein changes in one or more nucleotide bases does not affect the ability of the nucleic acid fragment to mediate alteration of gene expression by gene silencing through for example antisense or co-suppression technology. "Substantially similar" also refers to modifications of the nucleic acid fragments of the instant invention such as deletion or insertion of one or more nucleotides that do not substantially affect the functional properties of the resulting transcript vis-à-vis the ability to mediate gene silencing or alteration of the functional properties of the resulting protein molecule. It is therefore understood that the invention

5 encompasses more than the specific exemplary nucleotide or amino acid sequences and includes functional equivalents thereof.

10 Substantially similar nucleic acid fragments may be selected by screening nucleic acid fragments representing subfragments or modifications of the nucleic acid fragments of the instant invention, wherein one or more nucleotides are substituted, deleted and/or inserted, for their ability to affect the level of the polypeptide encoded by the unmodified nucleic acid fragment in a plant or plant cell. For example, a substantially similar nucleic acid fragment representing at least one of 30 contiguous nucleotides derived from the instant nucleic acid fragment can be constructed and introduced into a plant or plant cell. The level of the polypeptide encoded by the unmodified nucleic acid fragment present in a plant or plant cell exposed to the substantially similar nucleic acid fragment can then be compared to the level of the polypeptide in a plant or plant cell that is not exposed to the substantially similar nucleic acid fragment.

15 For example, it is well known in the art that antisense suppression and co-suppression of gene expression may be accomplished using nucleic acid fragments representing less than the entire coding region of a gene, and by nucleic acid fragments that do not share 100% sequence identity with the gene to be suppressed. Moreover, alterations in a nucleic acid fragment which result in the production of a chemically equivalent amino acid at a given site, but do not effect the functional properties of the encoded polypeptide, are well known in the art. Thus, a codon for the amino acid alanine, a hydrophobic amino acid, may be substituted by a codon encoding another less hydrophobic residue, such as glycine, or a more hydrophobic residue, such as valine, leucine, or isoleucine. Similarly, changes which result in substitution of one negatively charged residue for another, such as aspartic acid for glutamic acid, or one positively charged residue for another, such as lysine for arginine, can also be expected to produce a functionally equivalent product. Nucleotide changes which result in alteration of the N-terminal and C-terminal portions of the polypeptide molecule would also not be expected to alter the activity of the polypeptide. Each of the proposed modifications is well within the routine skill in the art, as is determination of retention of biological activity of the encoded products. Consequently, an isolated polynucleotide comprising a nucleotide sequence of at least one of 40 (preferably at least one of 30, most preferably at least one of 15) contiguous nucleotides derived from a nucleotide sequence selected from the group consisting of SEQ ID NOs:1, 3, 5, 7, and the complement of such nucleotide sequences may be used in methods of selecting an isolated polynucleotide that affects the expression of a polypeptide in a plant cell. A method of selecting an isolated polynucleotide that affects the level of expression of a polypeptide in a host cell (eukaryotic, such as plant or yeast, prokaryotic such as bacterial, or viral) may comprise the steps of: constructing an isolated polynucleotide of the present invention or an isolated chimeric gene of the present invention; introducing the isolated polynucleotide or the isolated chimeric

gene into a host cell; measuring the level a polypeptide in the host cell containing the isolated polynucleotide; and comparing the level of a polypeptide in the host cell containing the isolated polynucleotide with the level of a polypeptide in a host cell that does not contain the isolated polynucleotide.

Moreover, substantially similar nucleic acid fragments may also be characterized by their ability to hybridize. Estimates of such homology are provided by either DNA-DNA or DNA-RNA hybridization under conditions of stringency as is well understood by those skilled in the art (Hames and Higgins, Eds. (1985) *Nucleic Acid Hybridisation*, IRL Press, Oxford, U.K.). Stringency conditions can be adjusted to screen for moderately similar fragments, such as homologous sequences from distantly related organisms, to highly similar fragments, such as genes that duplicate functional enzymes from closely related organisms. Post-hybridization washes determine stringency conditions. One set of preferred conditions uses a series of washes starting with 6X SSC, 0.5% SDS at room temperature for 15 min, then repeated with 2X SSC, 0.5% SDS at 45°C for 30 min, and then repeated twice with 0.2X SSC, 0.5% SDS at 50°C for 30 min. A more preferred set of stringent conditions uses higher temperatures in which the washes are identical to those above except for the temperature of the final two 30 min washes in 0.2X SSC, 0.5% SDS was increased to 60°C. Another preferred set of highly stringent conditions uses two final washes in 0.1X SSC, 0.1% SDS at 65°C.

Substantially similar nucleic acid fragments of the instant invention may also be characterized by the percent identity of the amino acid sequences that they encode to the amino acid sequences disclosed herein, as determined by algorithms commonly employed by those skilled in this art. Suitable nucleic acid fragments (isolated polynucleotides of the present invention) encode polypeptides that are 80% identical to the amino acid sequences reported herein. Preferred nucleic acid fragments encode amino acid sequences that are 85% identical to the amino acid sequences reported herein. More preferred nucleic acid fragments encode amino acid sequences that are 90% identical to the amino acid sequences reported herein. Most preferred are nucleic acid fragments that encode amino acid sequences that are 95% identical to the amino acid sequences reported herein. Suitable nucleic acid fragments not only have the above homologies but typically encode a polypeptide having at least 50 amino acids, preferably 100 amino acids, more preferably 150 amino acids, still more preferably 200 amino acids, and most preferably 250 amino acids. Sequence alignments and percent identity calculations were performed using the Megalign program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Multiple alignment of the sequences was performed using the Clustal method of alignment (Higgins and Sharp (1989) *CABIOS*. 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise

alignments using the Clustal method were KTUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5.

A "substantial portion" of an amino acid or nucleotide sequence comprises an amino acid or a nucleotide sequence that is sufficient to afford putative identification of the protein or gene that the amino acid or nucleotide sequence comprises. Amino acid and nucleotide sequences can be evaluated either manually by one skilled in the art, or by using computer-based sequence comparison and identification tools that employ algorithms such as BLAST (Basic Local Alignment Search Tool; Altschul et al. (1993) *J. Mol. Biol.* 215:403-410; see also www.ncbi.nlm.nih.gov/BLAST/). In general, a sequence of ten or more contiguous amino acids or thirty or more contiguous nucleotides is necessary in order to putatively identify a polypeptide or nucleic acid sequence as homologous to a known protein or gene. Moreover, with respect to nucleotide sequences, gene-specific oligonucleotide probes comprising 30 or more contiguous nucleotides may be used in sequence-dependent methods of gene identification (e.g., Southern hybridization) and isolation (e.g., *in situ* hybridization of bacterial colonies or bacteriophage plaques). In addition, short oligonucleotides of 12 or more nucleotides may be used as amplification primers in PCR in order to obtain a particular nucleic acid fragment comprising the primers. Accordingly, a "substantial portion" of a nucleotide sequence comprises a nucleotide sequence that will afford specific identification and/or isolation of a nucleic acid fragment comprising the sequence. The instant specification teaches amino acid and nucleotide sequences encoding polypeptides that comprise one or more particular plant proteins. The skilled artisan, having the benefit of the sequences as reported herein, may now use all or a substantial portion of the disclosed sequences for purposes known to those skilled in this art. Accordingly, the instant invention comprises the complete sequences as reported in the accompanying Sequence Listing, as well as substantial portions of those sequences as defined above.

"Codon degeneracy" refers to divergence in the genetic code permitting variation of the nucleotide sequence without effecting the amino acid sequence of an encoded polypeptide. Accordingly, the instant invention relates to any nucleic acid fragment comprising a nucleotide sequence that encodes all or a substantial portion of the amino acid sequences set forth herein. The skilled artisan is well aware of the "codon-bias" exhibited by a specific host cell in usage of nucleotide codons to specify a given amino acid. Therefore, when synthesizing a nucleic acid fragment for improved expression in a host cell, it is desirable to design the nucleic acid fragment such that its frequency of codon usage approaches the frequency of preferred codon usage of the host cell.

"Synthetic nucleic acid fragments" can be assembled from oligonucleotide building blocks that are chemically synthesized using procedures known to those skilled in the art. These building blocks are ligated and annealed to form larger nucleic acid fragments which may then be enzymatically assembled to construct the entire desired nucleic acid fragment.

5 “Chemically synthesized”, as related to nucleic acid fragment, means that the component
nucleotides were assembled *in vitro*. Manual chemical synthesis of nucleic acid fragments
may be accomplished using well established procedures, or automated chemical synthesis
10 can be performed using one of a number of commercially available machines. Accordingly,
the nucleic acid fragments can be tailored for optimal gene expression based on optimization
5 of nucleotide sequence to reflect the codon bias of the host cell. The skilled artisan
appreciates the likelihood of successful gene expression if codon usage is biased towards
those codons favored by the host. Determination of preferred codons can be based on a
15 survey of genes derived from the host cell where sequence information is available.

10 “Gene” refers to a nucleic acid fragment that expresses a specific protein, including
regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding
sequences) the coding sequence. “Native gene” refers to a gene as found in nature with its
own regulatory sequences. “Chimeric gene” refers any gene that is not a native gene,
20 comprising regulatory and coding sequences that are not found together in nature.
Accordingly, a chimeric gene may comprise regulatory sequences and coding sequences that
15 are derived from different sources, or regulatory sequences and coding sequences derived
from the same source, but arranged in a manner different than that found in nature.
“Endogenous gene” refers to a native gene in its natural location in the genome of an
organism. A “foreign” gene refers to a gene not normally found in the host organism, but
20 that is introduced into the host organism by gene transfer. Foreign genes can comprise
native genes inserted into a non-native organism, or chimeric genes. A “transgene” is a gene
30 that has been introduced into the genome by a transformation procedure.

“Coding sequence” refers to a nucleotide sequence that codes for a specific amino acid
sequence. “Regulatory sequences” refer to nucleotide sequences located upstream (5' non-
35 coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence,
and which influence the transcription, RNA processing or stability, or translation of the
associated coding sequence. Regulatory sequences may include promoters, translation
leader sequences, introns, and polyadenylation recognition sequences.

40 “Promoter” refers to a nucleotide sequence capable of controlling the expression of a
coding sequence or functional RNA. In general, a coding sequence is located 3' to a
30 promoter sequence. The promoter sequence consists of proximal and more distal upstream
elements, the latter elements often referred to as enhancers. Accordingly, an “enhancer” is a
nucleotide sequence which can stimulate promoter activity and may be an innate element of
45 the promoter or a heterologous element inserted to enhance the level or tissue-specificity of
the promoter. Promoters may be derived in their entirety from a native gene, or be composed
35 of different elements derived from different promoters found in nature, or even comprise
synthetic nucleotide segments. It is understood by those skilled in the art that different
50 promoters may direct the expression of a gene in different tissues or cell types, or at

different stages of development, or in response to different environmental conditions. Promoters which cause a nucleic acid fragment to be expressed in most cell types at most times are commonly referred to as "constitutive promoters". New promoters of various types useful in plant cells are constantly being discovered; numerous examples may be found in the compilation by Okamuro and Goldberg (1989) *Biochemistry of Plants* 15:1-82. It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, nucleic acid fragments of different lengths may have identical promoter activity.

The "translation leader sequence" refers to a nucleotide sequence located between the promoter sequence of a gene and the coding sequence. The translation leader sequence is present in the fully processed mRNA upstream of the translation start sequence. The translation leader sequence may affect processing of the primary transcript to mRNA, mRNA stability or translation efficiency. Examples of translation leader sequences have been described (Turner and Foster (1995) *Mol. Biotechnol.* 3:225-236).

The "3' non-coding sequences" refer to nucleotide sequences located downstream of a coding sequence and include polyadenylation recognition sequences and other sequences encoding regulatory signals capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor. The use of different 3' non-coding sequences is exemplified by Ingelbrecht et al. (1989) *Plant Cell* 1:671-680.

"RNA transcript" refers to the product resulting from RNA polymerase-catalyzed transcription of a DNA sequence. When the RNA transcript is a perfect complementary copy of the DNA sequence, it is referred to as the primary transcript or it may be a RNA sequence derived from posttranscriptional processing of the primary transcript and is referred to as the mature RNA. "Messenger RNA (mRNA)" refers to the RNA that is without introns and that can be translated into polypeptide by the cell. "cDNA" refers to a double-stranded DNA that is complementary to and derived from mRNA. "Sense" RNA refers to an RNA transcript that includes the mRNA and so can be translated into a polypeptide by the cell. "Antisense RNA" refers to an RNA transcript that is complementary to all or part of a target primary transcript or mRNA and that blocks the expression of a target gene (see U.S. Patent No. 5,107,065, incorporated herein by reference). The complementarity of an antisense RNA may be with any part of the specific nucleotide sequence, i.e., at the 5' non-coding sequence, 3' non-coding sequence, introns, or the coding sequence. "Functional RNA" refers to sense RNA, antisense RNA, ribozyme RNA, or other RNA that may not be translated but yet has an effect on cellular processes.

The term "operably linked" refers to the association of two or more nucleic acid fragments on a single nucleic acid fragment so that the function of one is affected by the other. For example, a promoter is operably linked with a coding sequence when it is capable

5 of affecting the expression of that coding sequence (i.e., that the coding sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in sense or antisense orientation.

10 The term "expression", as used herein, refers to the transcription and stable accumulation of sense (mRNA) or antisense RNA derived from the nucleic acid fragment of the invention. Expression may also refer to translation of mRNA into a polypeptide. "Antisense inhibition" refers to the production of antisense RNA transcripts capable of suppressing the expression of the target protein. "Overexpression" refers to the production of a gene product in transgenic organisms that exceeds levels of production in normal or non-transformed organisms. "Co-suppression" refers to the production of sense RNA transcripts capable of suppressing the expression of identical or substantially similar foreign or endogenous genes (U.S. Patent No. 5,231,020, incorporated herein by reference).

20 "Altered levels" refers to the production of gene product(s) in transgenic organisms in amounts or proportions that differ from that of normal or non-transformed organisms.

25 "Mature" protein refers to a post-translationally processed polypeptide; i.e., one from which any pre- or propeptides present in the primary translation product have been removed. "Precursor" protein refers to the primary product of translation of mRNA; i.e., with pre- and propeptides still present. Pre- and propeptides may be but are not limited to intracellular localization signals.

30 A "chloroplast transit peptide" is an amino acid sequence which is translated in conjunction with a protein and directs the protein to the chloroplast or other plastid types present in the cell in which the protein is made. "Chloroplast transit sequence" refers to a nucleotide sequence that encodes a chloroplast transit peptide. A "signal peptide" is an amino acid sequence which is translated in conjunction with a protein and directs the protein to the secretory system (Chrispeels (1991) *Ann. Rev. Plant Phys. Plant Mol. Biol.* 42:21-53). If the protein is to be directed to a vacuole, a vacuolar targeting signal (*supra*) can further be added, or if to the endoplasmic reticulum, an endoplasmic reticulum retention signal (*supra*) may be added. If the protein is to be directed to the nucleus, any signal peptide present should be removed and instead a nuclear localization signal included (Raikhel (1992) *Plant Phys.* 100:1627-1632).

35 "Transformation" refers to the transfer of a nucleic acid fragment into the genome of a host organism, resulting in genetically stable inheritance. Host organisms containing the transformed nucleic acid fragments are referred to as "transgenic" organisms. Examples of methods of plant transformation include *Agrobacterium*-mediated transformation (De Blaere et al. (1987) *Meth. Enzymol.* 143:277) and particle-accelerated or "gene gun" transformation technology (Klein et al. (1987) *Nature (London)* 327:70-73; U.S. Patent No. 4,945,050, incorporated herein by reference).

Standard recombinant DNA and molecular cloning techniques used herein are well known in the art and are described more fully in Sambrook et al. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 1989 (hereinafter "Maniatis").

Nucleic acid fragments encoding at least a portion of several NPR1s have been isolated and identified by comparison of random plant cDNA sequences to public databases containing nucleotide and protein sequences using the BLAST algorithms well known to those skilled in the art. The nucleic acid fragments of the instant invention may be used to isolate cDNAs and genes encoding homologous proteins from the same or other plant species. Isolation of homologous genes using sequence-dependent protocols is well known in the art. Examples of sequence-dependent protocols include, but are not limited to, methods of nucleic acid hybridization, and methods of DNA and RNA amplification as exemplified by various uses of nucleic acid amplification technologies (e.g., polymerase chain reaction, ligase chain reaction).

For example, genes encoding other NPR1s, either as cDNAs or genomic DNAs, could be isolated directly by using all or a portion of the instant nucleic acid fragments as DNA hybridization probes to screen libraries from any desired plant employing methodology well known to those skilled in the art. Specific oligonucleotide probes based upon the instant nucleic acid sequences can be designed and synthesized by methods known in the art (Maniatis). Moreover, the entire sequences can be used directly to synthesize DNA probes by methods known to the skilled artisan such as random primer DNA labeling, nick translation, or end-labeling techniques, or RNA probes using available *in vitro* transcription systems. In addition, specific primers can be designed and used to amplify a part or all of the instant sequences. The resulting amplification products can be labeled directly during amplification reactions or labeled after amplification reactions, and used as probes to isolate full length cDNA or genomic fragments under conditions of appropriate stringency.

In addition, two short segments of the instant nucleic acid fragments may be used in polymerase chain reaction protocols to amplify longer nucleic acid fragments encoding homologous genes from DNA or RNA. The polymerase chain reaction may also be performed on a library of cloned nucleic acid fragments wherein the sequence of one primer is derived from the instant nucleic acid fragments, and the sequence of the other primer takes advantage of the presence of the polyadenylic acid tracts to the 3' end of the mRNA precursor encoding plant genes. Alternatively, the second primer sequence may be based upon sequences derived from the cloning vector. For example, the skilled artisan can follow the RACE protocol (Frohman et al. (1988) *Proc. Natl. Acad. Sci. USA* 85:8998-9002) to generate cDNAs by using PCR to amplify copies of the region between a single point in the transcript and the 3' or 5' end. Primers oriented in the 3' and 5' directions can be designed from the instant sequences. Using commercially available 3' RACE or 5' RACE systems

(BRL), specific 3' or 5' cDNA fragments can be isolated (Ohara et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:5673-5677; Loh et al. (1989) *Science* 243:217-220). Products generated by the 3' and 5' RACE procedures can be combined to generate full-length cDNAs (Frohman and Martin (1989) *Techniques* 1:165). Consequently, a polynucleotide comprising a nucleotide sequence of at least one of 40 (preferably one of at least 30, most preferably one of at least 15) contiguous nucleotides derived from a nucleotide sequence selected from the group consisting of SEQ ID NOs:1, 3, 5, 7, and the complement of such nucleotide sequences may be used in such methods to obtain a nucleic acid fragment encoding a substantial portion of an amino acid sequence of a polypeptide. The present invention relates to a method of obtaining a nucleic acid fragment encoding a substantial portion of a polypeptide of a gene (such as NPR1) preferably a substantial portion of a plant polypeptide of a gene, comprising the steps of: synthesizing an oligonucleotide primer comprising a nucleotide sequence of at least one of 40 (preferably at least one of 30, most preferably at least one of 15) contiguous nucleotides derived from a nucleotide sequence selected from the group consisting of SEQ ID NOs:1, 3, 5, 7, and the complement of such nucleotide sequences; and amplifying a nucleic acid fragment (preferably a cDNA inserted in a cloning vector) using the oligonucleotide primer. The amplified nucleic acid fragment preferably will encode a portion of a polypeptide.

Availability of the instant nucleotide and deduced amino acid sequences facilitates immunological screening of cDNA expression libraries. Synthetic peptides representing portions of the instant amino acid sequences may be synthesized. These peptides can be used to immunize animals to produce polyclonal or monoclonal antibodies with specificity for peptides or proteins comprising the amino acid sequences. These antibodies can be then be used to screen cDNA expression libraries to isolate full-length cDNA clones of interest (Lerner (1984) *Adv. Immunol.* 36:1-34; Maniatis).

The nucleic acid fragments of the instant invention may be used to create transgenic plants in which the disclosed polypeptides are present at higher levels than normal or in cell types or developmental stages in which they are not normally found. This would have the effect of altering the level of pathogen resistance in those cells. The NPR1 gene in *Arabidopsis thaliana* is involved in acquired pathogen resistance, thus its overexpression should allow the control of crop pathogens.

Overexpression of the proteins of the instant invention may be accomplished by first constructing a chimeric gene in which the coding region is operably linked to a promoter capable of directing expression of a gene in the desired tissues at the desired stage of development. For reasons of convenience, the chimeric gene may comprise promoter sequences and translation leader sequences derived from the same genes. 3' Non-coding sequences encoding transcription termination signals may also be provided. The instant chimeric gene may also comprise one or more introns in order to facilitate gene expression.

5 Plasmid vectors comprising the instant chimeric gene can then be constructed. The
choice of plasmid vector is dependent upon the method that will be used to transform host
plants. The skilled artisan is well aware of the genetic elements that must be present on the
10 plasmid vector in order to successfully transform, select and propagate host cells containing
the chimeric gene. The skilled artisan will also recognize that different independent
5 transformation events will result in different levels and patterns of expression (Jones et al.
(1985) *EMBO J.* 4:2411-2418; De Almeida et al. (1989) *Mol. Gen. Genetics* 218:78-86), and
15 thus that multiple events must be screened in order to obtain lines displaying the desired
expression level and pattern. Such screening may be accomplished by Southern analysis of
10 DNA, Northern analysis of mRNA expression, Western analysis of protein expression, or
phenotypic analysis.

20 For some applications it may be useful to direct the instant polypeptide to different
cellular compartments, or to facilitate its secretion from the cell. It is thus envisioned that
the chimeric gene described above may be further supplemented by altering the coding
15 sequence to encode the instant polypeptide with appropriate intracellular targeting sequences
such as transit sequences (Keegstra (1989) *Cell* 56:247-253), signal sequences or sequences
25 encoding endoplasmic reticulum localization (Chrispeels (1991) *Ann. Rev. Plant Phys. Plant*
Mol. Biol. 42:21-53), or nuclear localization signals (Raikhel (1992) *Plant*
Phys. 100:1627-1632) added and/or with targeting sequences that are already present
20 removed. While the references cited give examples of each of these, the list is not
exhaustive and more targeting signals of utility may be discovered in the future.

30 The instant polypeptide (or portions thereof) may be produced in heterologous host
cells, particularly in the cells of microbial hosts, and can be used to prepare antibodies to the
this protein by methods well known to those skilled in the art. The antibodies are useful for
35 detecting the polypeptide of the instant invention *in situ* in cells or *in vitro* in cell extracts.
Preferred heterologous host cells for production of the instant polypeptide are microbial
hosts. Microbial expression systems and expression vectors containing regulatory sequences
that direct high level expression of foreign proteins are well known to those skilled in the
40 art. Any of these could be used to construct a chimeric gene for production of the instant
polypeptide. This chimeric gene could then be introduced into appropriate microorganisms
30 via transformation to provide high level expression of the encoded NPR1. An example of a
vector for high level expression of the instant polypeptide in a bacterial host is provided
(Example 6).

45 All or a substantial portion of the nucleic acid fragments of the instant invention may
35 also be used as probes for genetically and physically mapping the genes that they are a part
of, and as markers for traits linked to those genes. Such information may be useful in plant
breeding in order to develop lines with desired phenotypes. For example, the instant nucleic
50 acid fragments may be used as restriction fragment length polymorphism (RFLP) markers.

5 Southern blots (Maniatis) of restriction-digested plant genomic DNA may be probed with
the nucleic acid fragments of the instant invention. The resulting banding patterns may then
be subjected to genetic analyses using computer programs such as MapMaker (Lander et al.
10 (1987) *Genomics* 1:174-181) in order to construct a genetic map. In addition, the nucleic
5 acid fragments of the instant invention may be used to probe Southern blots containing
restriction endonuclease-treated genomic DNAs of a set of individuals representing parent
and progeny of a defined genetic cross. Segregation of the DNA polymorphisms is noted
and used to calculate the position of the instant nucleic acid sequence in the genetic map
15 previously obtained using this population (Botstein et al. (1980) *Am. J. Hum. Genet.*
32:314-331).

10 The production and use of plant gene-derived probes for use in genetic mapping is
described in Bernatzky and Tanksley (1986) *Plant Mol. Biol. Reporter* 4:37-41. Numerous
20 publications describe genetic mapping of specific cDNA clones using the methodology
outlined above or variations thereof. For example, F2 intercross populations, backcross
15 populations, randomly mated populations, near isogenic lines, and other sets of individuals
may be used for mapping. Such methodologies are well known to those skilled in the art.

25 Nucleic acid probes derived from the instant nucleic acid sequences may also be used
for physical mapping (i.e., placement of sequences on physical maps; see Hoheisel et al. In:
Nonmammalian Genomic Analysis: A Practical Guide, Academic press 1996, pp. 319-346,
20 and references cited therein).

30 In another embodiment, nucleic acid probes derived from the instant nucleic acid
sequences may be used in direct fluorescence *in situ* hybridization (FISH) mapping (Trask
(1991) *Trends Genet.* 7:149-154). Although current methods of FISH mapping favor use of
large clones (several to several hundred KB; see Laan et al. (1995) *Genome Res.* 5:13-20),
35 improvements in sensitivity may allow performance of FISH mapping using shorter probes.

40 A variety of nucleic acid amplification-based methods of genetic and physical
mapping may be carried out using the instant nucleic acid sequences. Examples include
allele-specific amplification (Kazazian (1989) *J. Lab. Clin. Med.* 11:95-96), polymorphism
of PCR-amplified fragments (CAPS; Sheffield et al. (1993) *Genomics* 16:325-332), allele-
30 specific ligation (Landegren et al. (1988) *Science* 241:1077-1080), nucleotide extension
reactions (Sokolov (1990) *Nucleic Acid Res.* 18:3671), Radiation Hybrid Mapping (Walter
et al. (1997) *Nat. Genet.* 7:22-28) and Happy Mapping (Dear and Cook (1989) *Nucleic Acid*
45 *Res.* 17:6795-6807). For these methods, the sequence of a nucleic acid fragment is used to
design and produce primer pairs for use in the amplification reaction or in primer extension
35 reactions. The design of such primers is well known to those skilled in the art. In methods
employing PCR-based genetic mapping, it may be necessary to identify DNA sequence
differences between the parents of the mapping cross in the region corresponding to the
50

instant nucleic acid sequence. This, however, is generally not necessary for mapping methods.

EXAMPLES

The present invention is further defined in the following Examples, in which all parts and percentages are by weight and degrees are Celsius, unless otherwise stated. It should be understood that these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions.

EXAMPLE 1

Composition of cDNA Libraries; Isolation and Sequencing of cDNA Clones

cDNA libraries representing mRNAs from various corn, rice, and wheat tissues were prepared. The characteristics of the libraries are described below.

TABLE 2

cDNA Libraries from Corn, Rice, and Wheat

Library	Tissue	Clone
p0006	Corn Young Shoot	p0006.cbyvc82rx
rl0n	Rice 15 Day Old Leaf*	rl0n.pk0063.d10:fis
rr1	Rice Root of Two Week Old Developing Seedling	rr1.pk0001.a11:fis
wre1n	Wheat Root From 7 Day Old Etiolated Seedling*	wre1n.pk0122.c2:fis

* These libraries were normalized essentially as described in U.S. Patent No. 5,482,845, incorporated herein by reference.

cDNA libraries may be prepared by any one of many methods available. For example, the cDNAs may be introduced into plasmid vectors by first preparing the cDNA libraries in Uni-ZAP™ XR vectors according to the manufacturer's protocol (Stratagene Cloning Systems, La Jolla, CA). The Uni-ZAP™ XR libraries are converted into plasmid libraries according to the protocol provided by Stratagene. Upon conversion, cDNA inserts will be contained in the plasmid vector pBluescript. In addition, the cDNAs may be introduced directly into pre-cut Bluescript II SK(+) vectors (Stratagene) using T4 DNA ligase (New England Biolabs), followed by transfection into DH10B cells according to the manufacturer's protocol (GIBCO BRL Products). Once the cDNA inserts are in plasmid vectors, plasmid DNAs are prepared from randomly picked bacterial colonies containing recombinant pBluescript plasmids, or the insert cDNA sequences are amplified via polymerase chain reaction using primers specific for vector sequences flanking the inserted cDNA sequences. Amplified insert DNAs or plasmid DNAs are sequenced in dye-primer

sequencing reactions to generate partial cDNA sequences (expressed sequence tags or "ESTs"; see Adams et al., (1991) *Science* 252:1651-1656). The resulting ESTs are analyzed using a Perkin Elmer Model 377 fluorescent sequencer.

EXAMPLE 2

Identification of cDNA Clones

cDNA clones encoding NPR1s were identified by conducting BLAST (Basic Local Alignment Search Tool; Altschul et al. (1993) *J. Mol. Biol.* 215:403-410; see also www.ncbi.nlm.nih.gov/BLAST/) searches for similarity to sequences contained in the BLAST "nr" database (comprising all non-redundant GenBank CDS translations, sequences derived from the 3-dimensional structure Brookhaven Protein Data Bank, the last major release of the SWISS-PROT protein sequence database, EMBL, and DDBJ databases). The cDNA sequences obtained in Example 1 were analyzed for similarity to all publicly available DNA sequences contained in the "nr" database using the BLASTN algorithm provided by the National Center for Biotechnology Information (NCBI). The DNA sequences were translated in all reading frames and compared for similarity to all publicly available protein sequences contained in the "nr" database using the BLASTX algorithm (Gish and States (1993) *Nat. Genet.* 3:266-272) provided by the NCBI. For convenience, the P-value (probability) of observing a match of a cDNA sequence to a sequence contained in the searched databases merely by chance as calculated by BLAST are reported herein as "pLog" values, which represent the negative of the logarithm of the reported P-value. Accordingly, the greater the pLog value, the greater the likelihood that the cDNA sequence and the BLAST "hit" represent homologous proteins.

EXAMPLE 3

Characterization of cDNA Clones Encoding NPR1

The BLASTX search using the EST sequences from clones listed in Table 3 revealed similarity of the polypeptides encoded by the cDNAs to NPR1 from *Arabidopsis thaliana* (NCBI General Identifier No. 1773295). Shown in Table 3 are the BLAST results for individual ESTs ("EST"), the sequences of the entire cDNA inserts comprising the indicated cDNA clones ("FIS"), or sequences encoding the entire protein derived from an FIS, a contig, or an FIS and PCR ("CGS"):

TABLE 3
BLAST Results for Sequences Encoding Polypeptides Homologous to NPR1

Clone	Status	BLAST pLog Score 1773295
p0006.cbyvc82rx	EST	60.22
rl0n.pk0063.d10	CGS	138.00
rr1.pk0001.a11	FIS	91.22
wre1n.pk0122.c2	FIS	22.52

The data in Table 4 represents a calculation of the percent identity of the amino acid sequences set forth in SEQ ID NOs:2, 4, 6, and 8 and the *Arabidopsis thaliana* sequence (NCBI General Identifier No. 1773295).

TABLE 4

Percent Identity of Amino Acid Sequences Deduced From the Nucleotide Sequences of cDNA Clones Encoding Polypeptides Homologous to NPR1

SEQ ID NO.	Percent Identity to 1773295
2	40.3
4	42.4
6	39.1
8	33.7

Sequence alignments and percent identity calculations were performed using the Megalign program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Multiple alignment of the sequences was performed using the Clustal method of alignment (Higgins and Sharp (1989) *CABIOS*. 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments using the Clustal method were KTUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5. Sequence alignments and BLAST scores and probabilities indicate that the nucleic acid fragments comprising the instant cDNA clones encode a substantial portion of one corn, two rice, and one wheat NPR1. These sequences represent the first corn, rice, and wheat sequences encoding NPR1.

EXAMPLE 4

Expression of Chimeric Genes in Monocot Cells

A chimeric gene comprising a cDNA encoding the instant polypeptide in sense orientation with respect to the maize 27 kD zein promoter that is located 5' to the cDNA fragment, and the 10 kD zein 3' end that is located 3' to the cDNA fragment, can be constructed. The cDNA fragment of this gene may be generated by polymerase chain reaction (PCR) of the cDNA clone using appropriate oligonucleotide primers. Cloning sites (NcoI or SmaI) can be incorporated into the oligonucleotides to provide proper orientation of the DNA fragment when inserted into the digested vector pML103 as described below. Amplification is then performed in a standard PCR. The amplified DNA is then digested with restriction enzymes NcoI and SmaI and fractionated on an agarose gel. The appropriate band can be isolated from the gel and combined with a 4.9 kb NcoI-SmaI fragment of the plasmid pML103. Plasmid pML103 has been deposited under the terms of the Budapest Treaty at ATCC (American Type Culture Collection, 10801 University Blvd., Manassas,

5 VA 20110-2209), and bears accession number ATCC 97366. The DNA segment from
pML103 contains a 1.05 kb Sall-NcoI promoter fragment of the maize 27 kD zein gene and
10 a 0.96 kb SmaI-Sall fragment from the 3' end of the maize 10 kD zein gene in the vector
pGem9Zf(+) (Promega). Vector and insert DNA can be ligated at 15°C overnight,
5 essentially as described (Maniatis). The ligated DNA may then be used to transform *E. coli*
XL1-Blue (Epicurian Coli XL-1 Blue™; Stratagene). Bacterial transformants can be
screened by restriction enzyme digestion of plasmid DNA and limited nucleotide sequence
15 analysis using the dideoxy chain termination method (Sequenase™ DNA Sequencing Kit;
U.S. Biochemical). The resulting plasmid construct would comprise a chimeric gene
10 encoding, in the 5' to 3' direction, the maize 27 kD zein promoter, a cDNA fragment
encoding the instant polypeptide, and the 10 kD zein 3' region.

20 The chimeric gene described above can then be introduced into corn cells by the
following procedure. Immature corn embryos can be dissected from developing caryopses
derived from crosses of the inbred corn lines H99 and LH132. The embryos are isolated 10
15 to 11 days after pollination when they are 1.0 to 1.5 mm long. The embryos are then placed
with the axis-side facing down and in contact with agarose-solidified N6 medium (Chu et al.
25 (1975) *Sci. Sin. Peking* 18:659-668). The embryos are kept in the dark at 27°C. Friable
embryogenic callus consisting of undifferentiated masses of cells with somatic
proembryoids and embryoids borne on suspensor structures proliferates from the scutellum
20 of these immature embryos. The embryogenic callus isolated from the primary explant can
be cultured on N6 medium and sub-cultured on this medium every 2 to 3 weeks.

30 The plasmid, p35S/Ac (obtained from Dr. Peter Eckes, Hoechst Ag, Frankfurt,
Germany) may be used in transformation experiments in order to provide for a selectable
marker. This plasmid contains the *Pat* gene (see European Patent Publication 0 242 236)
35 25 which encodes phosphinothricin acetyl transferase (PAT). The enzyme PAT confers
resistance to herbicidal glutamine synthetase inhibitors such as phosphinothricin. The *pat*
gene in p35S/Ac is under the control of the 35S promoter from Cauliflower Mosaic Virus
(Odell et al. (1985) *Nature* 313:810-812) and the 3' region of the nopaline synthase gene
40 from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*.

30 The particle bombardment method (Klein et al. (1987) *Nature* 327:70-73) may be used
to transfer genes to the callus culture cells. According to this method, gold particles (1 µm
in diameter) are coated with DNA using the following technique. Ten µg of plasmid DNAs
45 are added to 50 µL of a suspension of gold particles (60 mg per mL). Calcium chloride
(50 µL of a 2.5 M solution) and spermidine free base (20 µL of a 1.0 M solution) are added
35 to the particles. The suspension is vortexed during the addition of these solutions. After
10 minutes, the tubes are briefly centrifuged (5 sec at 15,000 rpm) and the supernatant
50 removed. The particles are resuspended in 200 µL of absolute ethanol, centrifuged again
and the supernatant removed. The ethanol rinse is performed again and the particles

5 resuspended in a final volume of 30 μ L of ethanol. An aliquot (5 μ L) of the DNA-coated
gold particles can be placed in the center of a Kapton™ flying disc (Bio-Rad Labs). The
particles are then accelerated into the corn tissue with a Biolistic™ PDS-1000/He (Bio-Rad
10 Instruments, Hercules CA), using a helium pressure of 1000 psi, a gap distance of 0.5 cm
5 and a flying distance of 1.0 cm.

For bombardment, the embryogenic tissue is placed on filter paper over agarose-
solidified N6 medium. The tissue is arranged as a thin lawn and covered a circular area of
15 about 5 cm in diameter. The petri dish containing the tissue can be placed in the chamber of
the PDS-1000/He approximately 8 cm from the stopping screen. The air in the chamber is
10 then evacuated to a vacuum of 28 inches of Hg. The macrocarrier is accelerated with a
helium shock wave using a rupture membrane that bursts when the He pressure in the shock
tube reaches 1000 psi.

20 Seven days after bombardment the tissue can be transferred to N6 medium that
contains glufosinate (2 mg per liter) and lacks casein or proline. The tissue continues to
15 grow slowly on this medium. After an additional 2 weeks the tissue can be transferred to
fresh N6 medium containing glufosinate. After 6 weeks, areas of about 1 cm in diameter
25 of actively growing callus can be identified on some of the plates containing the glufosinate-
supplemented medium. These calli may continue to grow when sub-cultured on the
selective medium.

20 Plants can be regenerated from the transgenic callus by first transferring clusters of
tissue to N6 medium supplemented with 0.2 mg per liter of 2,4-D. After two weeks the
30 tissue can be transferred to regeneration medium (Fromm et al. (1990) *Bio/Technology*
8:833-839).

EXAMPLE 5

Expression of Chimeric Genes in Dicot Cells

35 A seed-specific expression cassette composed of the promoter and transcription
terminator from the gene encoding the β subunit of the seed storage protein phaseolin from
the bean *Phaseolus vulgaris* (Doyle et al. (1986) *J. Biol. Chem.* 261:9228-9238) can be used
40 for expression of the instant polypeptide in transformed soybean. The phaseolin cassette
30 includes about 500 nucleotides upstream (5') from the translation initiation codon and about
1650 nucleotides downstream (3') from the translation stop codon of phaseolin. Between the
5' and 3' regions are the unique restriction endonuclease sites Nco I (which includes the ATG
45 translation initiation codon), Sma I, Kpn I and Xba I. The entire cassette is flanked by
Hind III sites.

35 The cDNA fragment of this gene may be generated by polymerase chain reaction
(PCR) of the cDNA clone using appropriate oligonucleotide primers. Cloning sites can be
50 incorporated into the oligonucleotides to provide proper orientation of the DNA fragment
when inserted into the expression vector. Amplification is then performed as described

5 above, and the isolated fragment is inserted into a pUC18 vector carrying the seed expression cassette.

10 Soybean embryos may then be transformed with the expression vector comprising sequences encoding the instant polypeptide. To induce somatic embryos, cotyledons, 3-5 mm in length dissected from surface sterilized, immature seeds of the soybean cultivar A2872, can be cultured in the light or dark at 26°C on an appropriate agar medium for 6-10 weeks. Somatic embryos which produce secondary embryos are then excised and placed into a suitable liquid medium. After repeated selection for clusters of somatic embryos which multiplied as early, globular staged embryos, the suspensions are maintained as described below.

15 Soybean embryogenic suspension cultures can maintained in 35 mL liquid media on a rotary shaker, 150 rpm, at 26°C with florescent lights on a 16:8 hour day/night schedule. Cultures are subcultured every two weeks by inoculating approximately 35 mg of tissue into 35 mL of liquid medium.

20 Soybean embryogenic suspension cultures may then be transformed by the method of particle gun bombardment (Klein et al. (1987) *Nature* (London) 327:70-73, U.S. Patent No. 4,945,050). A DuPont Biolistic™ PDS1000/HE instrument (helium retrofit) can be used for these transformations.

25 A selectable marker gene which can be used to facilitate soybean transformation is a chimeric gene composed of the 35S promoter from Cauliflower Mosaic Virus (Odell et al. (1985) *Nature* 313:810-812), the hygromycin phosphotransferase gene from plasmid pJR225 (from *E. coli*; Gritz et al. (1983) *Gene* 25:179-188) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*. The seed expression cassette comprising the phaseolin 5' region, the fragment encoding the instant polypeptide and the phaseolin 3' region can be isolated as a restriction fragment. This fragment can then be inserted into a unique restriction site of the vector carrying the marker gene.

30 To 50 µL of a 60 mg/mL 1 µm gold particle suspension is added (in order): 5 µL DNA (1 µg/µL), 20 µl spermidine (0.1 M), and 50 µL CaCl₂ (2.5 M). The particle preparation is then agitated for three minutes, spun in a microfuge for 10 seconds and the supernatant removed. The DNA-coated particles are then washed once in 400 µL 70% ethanol and resuspended in 40 µL of anhydrous ethanol. The DNA/particle suspension can be sonicated three times for one second each. Five µL of the DNA-coated gold particles are then loaded on each macro carrier disk.

35 Approximately 300-400 mg of a two-week-old suspension culture is placed in an empty 60x15 mm petri dish and the residual liquid removed from the tissue with a pipette. For each transformation experiment, approximately 5-10 plates of tissue are normally bombarded. Membrane rupture pressure is set at 1100 psi and the chamber is evacuated to a vacuum of 28 inches mercury. The tissue is placed approximately 3.5 inches away from the

5 retaining screen and bombarded three times. Following bombardment, the tissue can be divided in half and placed back into liquid and cultured as described above.

10 Five to seven days post bombardment, the liquid media may be exchanged with fresh media, and eleven to twelve days post bombardment with fresh media containing 50 mg/mL
5 hygromycin. This selective media can be refreshed weekly. Seven to eight weeks post bombardment, green, transformed tissue may be observed growing from untransformed, necrotic embryogenic clusters. Isolated green tissue is removed and inoculated into
15 individual flasks to generate new, clonally propagated, transformed embryogenic suspension cultures. Each new line may be treated as an independent transformation event. These
20 suspensions can then be subcultured and maintained as clusters of immature embryos or regenerated into whole plants by maturation and germination of individual somatic embryos.

EXAMPLE 6

Expression of Chimeric Genes in Microbial Cells

20 The cDNAs encoding the instant polypeptide can be inserted into the T7 *E. coli* expression vector pBT430. This vector is a derivative of pET-3a (Rosenberg et al. (1987)
15 *Gene* 56:125-135) which employs the bacteriophage T7 RNA polymerase/T7 promoter system. Plasmid pBT430 was constructed by first destroying the EcoR I and Hind III sites in
25 pET-3a at their original positions. An oligonucleotide adaptor containing EcoR I and Hind III sites was inserted at the BamH I site of pET-3a. This created pET-3aM with
30 additional unique cloning sites for insertion of genes into the expression vector. Then, the Nde I site at the position of translation initiation was converted to an Nco I site using
oligonucleotide-directed mutagenesis. The DNA sequence of pET-3aM in this region,
5'-CATATGG, was converted to 5'-CCCATGG in pBT430.

35 Plasmid DNA containing a cDNA may be appropriately digested to release a nucleic acid fragment encoding the protein. This fragment may then be purified on a 1% NuSieve
25 GTG™ low melting agarose gel (FMC). Buffer and agarose contain 10 µg/ml ethidium bromide for visualization of the DNA fragment. The fragment can then be purified from the
40 agarose gel by digestion with GELase™ (Epicentre Technologies) according to the manufacturer's instructions, ethanol precipitated, dried and resuspended in 20 µL of water.
30 Appropriate oligonucleotide adapters may be ligated to the fragment using T4 DNA ligase (New England Biolabs, Beverly, MA). The fragment containing the ligated adapters can be
45 purified from the excess adapters using low melting agarose as described above. The vector pBT430 is digested, dephosphorylated with alkaline phosphatase (NEB) and deproteinized
with phenol/chloroform as described above. The prepared vector pBT430 and fragment can
35 then be ligated at 16°C for 15 hours followed by transformation into DH5 electrocompetent cells (GIBCO BRL). Transformants can be selected on agar plates containing LB media and
50 100 µg/mL ampicillin. Transformants containing the gene encoding the instant polypeptide

5 are then screened for the correct orientation with respect to the T7 promoter by restriction enzyme analysis.

10 For high level expression, a plasmid clone with the cDNA insert in the correct orientation relative to the T7 promoter can be transformed into *E. coli* strain BL21(DE3) (Studier et al. (1986) *J. Mol. Biol.* 189:113-130). Cultures are grown in LB medium containing ampicillin (100 mg/L) at 25°C. At an optical density at 600 nm of approximately 1, IPTG (isopropylthio- β -galactoside, the inducer) can be added to a final concentration of 0.4 mM and incubation can be continued for 3 h at 25°. Cells are then harvested by centrifugation and re-suspended in 50 μ L of 50 mM Tris-HCl at pH 8.0 containing 0.1 mM DTT and 0.2 mM phenyl methylsulfonyl fluoride. A small amount of 1 mm glass beads can be added and the mixture sonicated 3 times for about 5 seconds each time with a microprobe sonicator. The mixture is centrifuged and the protein concentration of the supernatant determined. One μ g of protein from the soluble fraction of the culture can be separated by SDS-polyacrylamide gel electrophoresis. Gels can be observed for protein bands migrating at the expected molecular weight.

Claims

5

10

15

20

25

30

35

40

45

50

55

CLAIMS

What is claimed is:

1. A composition consisting of an isolated polynucleotide comprising a nucleotide sequence encoding a first polypeptide of at least 80 amino acids that has at least 75% identity based on the Clustal method of alignment when compared to a polypeptide selected from the group consisting of an NPR1 polypeptide of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, and SEQ ID NO:8, or an isolated polynucleotide comprising the complement of the nucleotide sequence.
2. The composition of Claim 1, wherein the isolated nucleotide sequence consists of a nucleic acid sequence selected from the group consisting of SEQ ID NOs:1, 3, 5, and 7 that codes for the polypeptide selected from the group consisting of SEQ ID NOs:2, 4, 6, and 8.
3. The composition of Claim 1 wherein the isolated polynucleotide is DNA.
4. The composition of Claim 1 wherein the isolated polynucleotide is RNA.
5. A chimeric gene comprising the isolated polynucleotide of Claim 1 operably linked to suitable regulatory sequences.
6. An isolated host cell comprising the chimeric gene of Claim 5.
7. An isolated host cell comprising an isolated polynucleotide of Claim 1.
8. The isolated host cell of Claim 7 wherein the isolated host cell is selected from the group consisting of yeast, bacteria, plant, and virus.
9. A virus comprising the isolated polynucleotide of Claim 1.
10. A composition consisting of a polypeptide of at least 80 amino acids that has at least 75% identity based on the Clustal method of alignment when compared to a polypeptide selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, and SEQ ID NO:8.
11. A method of selecting an isolated polynucleotide that affects the level of expression of an NPR1 polypeptide in a plant cell, the method comprising the steps of:
 - (a) constructing an isolated polynucleotide comprising a nucleotide sequence of at least one of 30 contiguous nucleotides derived from a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, 3, 5, 7, and the complement of such nucleotide sequences;
 - (b) introducing the isolated polynucleotide into a plant cell; and
 - (c) measuring the level of a polypeptide in the plant cell containing the polynucleotide.
12. The method of Claim 11 wherein the isolated polynucleotide consists of a nucleotide sequence selected from the group consisting of SEQ ID NOs:1, 3, 5, and 7 that codes for the polypeptide selected from the group consisting of SEQ ID NOs:2, 4, 6, and 8.

5
13. A method of selecting an isolated polynucleotide that affects the level of expression of an NPR1 polypeptide in a plant cell, the method comprising the steps of:

- 10
5
(a) constructing an isolated polynucleotide of Claim 1;
(b) introducing the isolated polynucleotide into a plant cell; and
(c) measuring the level of polypeptide in the plant cell containing the polynucleotide.

14. A method of obtaining a nucleic acid fragment encoding an NPR1 polypeptide comprising the steps of:

- 15
10
(a) synthesizing an oligonucleotide primer comprising a nucleotide sequence of at least one of 30 contiguous nucleotides derived from a nucleotide sequence selected from the group consisting of SEQ ID NOs:1, 3, 5, 7 and the complement of such nucleotide sequences; and

20
(b) amplifying a nucleic acid sequence using the oligonucleotide primer.

15
15. A method of obtaining a nucleic acid fragment encoding the amino acid sequence encoding an NPR1 polypeptide comprising the steps of:

25
(a) probing a cDNA or genomic library with an isolated polynucleotide comprising a nucleotide sequence of at least one of 30 contiguous nucleotides derived from a nucleotide sequence selected from the group consisting of SEQ ID NOs:1, 3, 5, 7, and the complement of such nucleotide sequences;

20
(b) identifying a DNA clone that hybridizes with the isolated polynucleotide;

30
(c) isolating the identified DNA clone; and

(d) sequencing the cDNA or genomic fragment that comprises the isolated DNA clone.

35
25 16. An expression cassette comprising at least one nucleic acid of Claim 1 operably linked to a promoter.

17. A method for positive selection of a transformed cell comprising:

(a) transforming a plant cell with the expression cassette of Claim 5; and

40
30 (b) growing the transformed plant cell under conditions allowing expression of the polynucleotide in an amount sufficient to induce disease resistance to provide a positive selection means.

18. The method of Claim 17, wherein the plant cell is a monocot.

SEQUENCE LISTING

<110> E. I. du Pont de Nemours and Company

<120> Disease Resistance Factors

<130> BB1252

<140>

<141>

<150> 60/107,242

<151> 1998-11-05

<160> 8

<170> Microsoft Office 97

<210> 1

<211> 1227

<212> DNA

<213> Zea mays

```

<400> 1
agttgatgag ataaaaaatt tgcgcaagaa gtcacaaaact gctgatgggtg atacgttcac 60
ttcggaccct gtgcatgaga aaagagtcag aagaatccac agggcacttg actctgatga 120
tggtgagctt gtgaagttgc ttottaatga gtccgacatc acattagatg atgccaacgc 180
attacactat gctgcttctt actgtgatcc taaagttgtc tcagagctgt tagatttggc 240
aatggctaac ttaaatttga agaatagcgc tgggtacaca gcactccact tggctgctat 300
gaggagagaa ccagctataa tcatgtgtct ccttaacaaa ggggcacaaatg tgtcacaact 360
gacagctgat ggcaggagcg caattggtat ttgtcggagg ttaacaagag caaaagacta 420
caatacaaaag atggagcagg gtcaagaatc aaataaagat aggtctgtga tagatattct 480
agagagggag atgatgcgga atcctatggc ggtggaagat gccgtcacct cgcctttgtt 540
ggcagatgat cttcacatga agcttctcta cctggaaaac agagttgcat ttgctagatt 600
gttctttcct gctgaagcca aggtcgccat gcaaatcgca caagcagaca ccacagaaga 660
attcggcggt atagttgcag ttgcagcaag cacttctggt aaactgaggg aggtggacct 720
taatgagacg ccagtgacac aaaacaaaag gctccgttca aggttagatg cactgatgaa 780
aacagtggag ctgggccgtc ggtacttccc gaactgctcg caggtgctgg acaagttcct 840
ggaggacgat ctgcgggaag gtctggacca gttctacctc cagaggggca cagccgatga 900
gcagaagggt aagaggatgc gcttctgcga gctgaaagag gacgtgctga aggcgtttag 960
caaggacaag gcggagggca gcgtgttctc gggcctgtcc tcgtcgtcgt cgtgctcgcc 1020
gccccagaag tatgccaga ggtgatcaag gcaccagttt ttgccgtata gtttgttatc 1080
atggtcttcg agacttggac ccggacagca tatagggaca tgtacacctg tgtatgtata 1140
gtgcttacaa ttggcgtaag tagaactata tgtatggaac ataaggaaac atggcaggaa 1200
caccgtgcaa aaagatgaaa aaaaaaa 1227

```

<210> 2

<211> 325

<212> PRT

<213> Zea mays

<400> 2

```

Pro Val His Glu Lys Arg Val Arg Arg Ile His Arg Ala Leu Asp Ser
  1             5             10             15

```

```

Asp Asp Val Glu Leu Val Lys Leu Leu Leu Asn Glu Ser Asp Ile Thr
          20             25             30

```

```

Leu Asp Asp Ala Asn Ala Leu His Tyr Ala Ala Ser Tyr Cys Asp Pro
          35             40             45

```

Lys Val Val Ser Glu Leu Leu Asp Leu Ala Met Ala Asn Leu Asn Leu
 50 55 60
 Lys Asn Ser Arg Gly Tyr Thr Ala Leu His Leu Ala Ala Met Arg Arg
 65 70 75 80
 Glu Pro Ala Ile Ile Met Cys Leu Leu Asn Lys Gly Ala Asn Val Ser
 85 90 95
 Gln Leu Thr Ala Asp Gly Arg Ser Ala Ile Gly Ile Cys Arg Arg Leu
 100 105 110
 Thr Arg Ala Lys Asp Tyr Asn Thr Lys Met Glu Gln Gly Gln Glu Ser
 115 120 125
 Asn Lys Asp Arg Leu Cys Ile Asp Ile Leu Glu Arg Glu Met Met Arg
 130 135 140
 Asn Pro Met Ala Val Glu Asp Ala Val Thr Ser Pro Leu Leu Ala Asp
 145 150 155 160
 Asp Leu His Met Lys Leu Leu Tyr Leu Glu Asn Arg Val Ala Phe Ala
 165 170 175
 Arg Leu Phe Phe Pro Ala Glu Ala Lys Val Ala Met Gln Ile Ala Gln
 180 185 190
 Ala Asp Thr Thr Glu Glu Phe Gly Gly Ile Val Ala Val Ala Ala Ser
 195 200 205
 Thr Ser Gly Lys Leu Arg Glu Val Asp Leu Asn Glu Thr Pro Val Thr
 210 215 220
 Gln Asn Lys Arg Leu Arg Ser Arg Val Asp Ala Leu Met Lys Thr Val
 225 230 235 240
 Glu Leu Gly Arg Arg Tyr Phe Pro Asn Cys Ser Gln Val Leu Asp Lys
 245 250 255
 Phe Leu Glu Asp Asp Leu Pro Glu Gly Leu Asp Gln Phe Tyr Leu Gln
 260 265 270
 Arg Gly Thr Ala Asp Glu Gln Lys Val Lys Arg Met Arg Phe Cys Glu
 275 280 285
 Leu Lys Glu Asp Val Leu Lys Ala Phe Ser Lys Asp Lys Ala Glu Gly
 290 295 300
 Ser Val Phe Ser Gly Leu Ser Ser Ser Ser Ser Cys Ser Pro Pro Gln
 305 310 315 320
 Lys Tyr Ala Gln Arg
 325

<210> 3
 <211> 2194
 <212> DNA
 <213> Oryza sativa

```

<400> 3
cccccgggct gcaggaattc ggcacgaggc tggggcggga ggctctctcc tcgctcgc 60
tcgccacgcc ggcgcgcgac ggcacgcgcc gtggtcagct ggtcgcgggt gcgggtgagg 120
gtgcgcgaatg gagccgcga ccagccacgt caccacgcg ttctccgact cggacagcgc 180
gtccgtggag gaggggggcg ccgacgcgga cgcgcagctg gagggcgtcc gccgcctctc 240
cgacaacctc gccgcggcgt tccgctcgcc cgaggacttc gcgttctcgc ccgacgcgcg 300
catcgccgtc ccggggcgcg ggcggcgcg cgccgacctg ctgggtgcacc gctgcgtgct 360
ctccgcgcgg agccctctcc tgcgcggcgt cttcgccgc ccgcgcgcgc ccgccgcagg 420
cggcgccggc gaggatggcg gcgagaggct ggagctcccg gaactcctcg gcggcgccgg 480
cgaggagggt gaggtcgggt acgaggcgct gcggctggtg ctgcactacc tctacagcgc 540
ccgcgtcggc gacctgcca aggcgcgcgt cctctgcgtc gacgaggact gcgccacgt 600
cgggtgccac ccgcgcgtcg cgttcattgg gcaggctctc ttccgcgcct ccaccttcca 660
ggtcgcgcag ctaccaaac tcttcagcgc gcgtctctct gatgtccttg ataagggtga 720
ggtagataac cttctattga tcttatctgt tggcaactta tgcaacaaat cttgcatgaa 780
actgcttgaa agatgccttg atatggtagt ccggtcaaac cttgacatga ttactcttga 840
gaagtcatgt cctccagatg ttatcaagca gattattgat gcacgcctaa gcctcggatt 900
aatttcacca gaaaacaagg gatttcctaa caaacatgtg aggaggatac acagagccct 960
tgactctgac gatgtagagc tagtcaggat gctgctcact gaaggacaga caaatcttga 1020
tgatgcgttt gcactgcact acgcgcgtcg acattgtgac tccaaaatta caaccgagct 1080
tttgatctc gcacttgca atgttaatca tagaaaccca agagggtata ctgttcttca 1140
cattgctcgc aggcgaagag agcctaaaat cattgtctcc ctttaacca agggggcctc 1200
gccagcagat gttacattcg atgggagaaa agcgggtcaa atctcaaaaa gactaacaaa 1260
acaaggggat tactttggg ttaccgaaga aggaaaacct tctccaaaag ataggttatg 1320
tattgaaata ctggagcaag ctgaaagaag ggaccacaa ctccggagaag catcagtttc 1380
tcttgcaatg gcagggtgag gtctacgagg aaggttgctg tatcttgaaa accgagttgc 1440
tttggcgagg attatgttcc cgatggagcc aagagttagc atggatattg ctcaagtgga 1500
tggaactttg gaatttaacc tgggttctcg tgcaaatcca cctcctgaaa gacaacggac 1560
aactgttgat ctaaatgaaa gtcccttcat aatgaaagaa gaacacttag ctccgatgac 1620
ggcactctcc aaaacagtgg agctcgggaa acgctttttc ccgcgatgtt cgaacgtgct 1680
cgacaagatc atggatgatg aaactgatcc gggttccctc ggaagagaca cgtccgcgga 1740
gaagaggaag aggtttcatg acctgcagga tgttcttcag aaggcattcc acgaggacaa 1800
ggaggagaat gacaggtcgg ggcctctcgc gtctcgtcgc tcgacatcga tcggggccat 1860
tcgaccaagg agatgaacac cattgctccc aaatagttgc catattgata gctaactgtc 1920
ctcctggagc tactcacctg atgggttgcct tctgtcaatt gccccccaaa tatattctca 1980
atgggttagg cttgtacagt attagttcct acagctattg ccccgccaat tgtgaaacgc 2040
agaagtttca ctagtgtctg tactcgaggt gtaatacaag tgcttgaatt ttgagttgta 2100
cttggaaatt ccagtgtttt gctcgtaaaa atgagatgat ttcttggtc ccaaaaaaaa 2160
aaaaaaaa aactcgaggg ggggcccggg accc

```

```

<210> 4
<211> 578
<212> PRT
<213> Oryza sativa

```

```

<400> 4
Thr Ser His Val Thr Asn Ala Phe Ser Asp Ser Asp Ser Ala Ser Val
1 5 10 15

Glu Glu Gly Gly Ala Asp Ala Asp Ala Asp Val Glu Ala Leu Arg Arg
20 25 30

Leu Ser Asp Asn Leu Ala Ala Ala Phe Arg Ser Pro Glu Asp Phe Ala
35 40 45

Phe Leu Ala Asp Ala Arg Ile Ala Val Pro Gly Gly Gly Gly Gly
50 55 60

Gly Asp Leu Leu Val His Arg Cys Val Leu Ser Ala Arg Ser Pro Phe
65 70 75 80

```


Leu Arg Gly Val Phe Ala Arg Arg Ala Ala Ala Ala Ala Gly Gly Gly
 85 90 95
 Gly Glu Asp Gly Gly Glu Arg Leu Glu Leu Arg Glu Leu Leu Gly Gly
 100 105 110
 Gly Gly Glu Glu Val Glu Val Gly Tyr Glu Ala Leu Arg Leu Val Leu
 115 120 125
 Asp Tyr Leu Tyr Ser Gly Arg Val Gly Asp Leu Pro Lys Ala Ala Cys
 130 135 140
 Leu Cys Val Asp Glu Asp Cys Ala His Val Gly Cys His Pro Ala Val
 145 150 155 160
 Ala Phe Met Ala Gln Val Leu Phe Ala Ala Ser Thr Phe Gln Val Ala
 165 170 175
 Glu Leu Thr Asn Leu Phe Gln Arg Arg Leu Leu Asp Val Leu Asp Lys
 180 185 190
 Val Glu Val Asp Asn Leu Leu Leu Ile Leu Ser Val Ala Asn Leu Cys
 195 200 205
 Asn Lys Ser Cys Met Lys Leu Leu Glu Arg Cys Leu Asp Met Val Val
 210 215 220
 Arg Ser Asn Leu Asp Met Ile Thr Leu Glu Lys Ser Leu Pro Pro Asp
 225 230 235 240
 Val Ile Lys Gln Ile Ile Asp Ala Arg Leu Ser Leu Gly Leu Ile Ser
 245 250 255
 Pro Glu Asn Lys Gly Phe Pro Asn Lys His Val Arg Arg Ile His Arg
 260 265 270
 Ala Leu Asp Ser Asp Asp Val Glu Leu Val Arg Met Leu Leu Thr Glu
 275 280 285
 Gly Gln Thr Asn Leu Asp Asp Ala Phe Ala Leu His Tyr Ala Val Glu
 290 295 300
 His Cys Asp Ser Lys Ile Thr Thr Glu Leu Leu Asp Leu Ala Leu Ala
 305 310 315 320
 Asp Val Asn His Arg Asn Pro Arg Gly Tyr Thr Val Leu His Ile Ala
 325 330 335
 Ala Arg Arg Arg Glu Pro Lys Ile Ile Val Ser Leu Leu Thr Lys Gly
 340 345 350
 Ala Arg Pro Ala Asp Val Thr Phe Asp Gly Arg Lys Ala Val Gln Ile
 355 360 365
 Ser Lys Arg Leu Thr Lys Gln Gly Asp Tyr Phe Gly Val Thr Glu Glu
 370 375 380
 Gly Lys Pro Ser Pro Lys Asp Arg Leu Cys Ile Glu Ile Leu Glu Gln
 385 390 395 400

Ala Glu Arg Arg Asp Pro Gln Leu Gly Glu Ala Ser Val Ser Leu Ala
 405 410 415

Met Ala Gly Glu Ser Leu Arg Gly Arg Leu Leu Tyr Leu Glu Asn Arg
 420 425 430

Val Ala Leu Ala Arg Ile Met Phe Pro Met Glu Ala Arg Val Ala Met
 435 440 445

Asp Ile Ala Gln Val Asp Gly Thr Leu Glu Phe Asn Leu Gly Ser Gly
 450 455 460

Ala Asn Pro Pro Pro Glu Arg Gln Arg Thr Thr Val Asp Leu Asn Glu
 465 470 475 480

Ser Pro Phe Ile Met Lys Glu Glu His Leu Ala Arg Met Thr Ala Leu
 485 490 495

Ser Lys Thr Val Glu Leu Gly Lys Arg Phe Phe Pro Arg Cys Ser Asn
 500 505 510

Val Leu Asp Lys Ile Met Asp Asp Glu Thr Asp Pro Val Ser Leu Gly
 515 520 525

Arg Asp Thr Ser Ala Glu Lys Arg Lys Arg Phe His Asp Leu Gln Asp
 530 535 540

Val Leu Gln Lys Ala Phe His Glu Asp Lys Glu Glu Asn Asp Arg Ser
 545 550 555 560

Gly Leu Ser Ser Ser Ser Ser Ser Thr Ser Ile Gly Ala Ile Arg Pro
 565 570 575

Arg Arg

<210> 5
 <211> 2069
 <212> DNA
 <213> Oryza sativa

<220>
 <221> unsure
 <222> (65)

<400> 5
 gttgtrtggga attgtgagcg ataacaattt macacaggaa acagctatga ccatgattac 60
 gccangcgm g caattaaccs tcactaaagg gaacaaaagg tggagcwcca ccgcggtggc 120
 ggccgctcta gaavtagtgg atcccccggg ctgcaggaat tcggcacgag gcgcgatgcc 180
 ttccctgtgc tgcctgggtta cctgtacacg ggcaagctcc ggccggcgcc ggatgacgtg 240
 gtgtcctgcg ccgaccccat gtgcccgcac gactcgtgcc cgccggcgat caggttcaac 300
 gtcgagcaaa tgtacgcggc gtgggcgttc aagatcaccg agctcatctc gctgttccag 360
 cgacggcttc ttaacttcgt cgataagact ctagtagaag atgttcttcc aattctgcaa 420
 gttgcttttc attcagagct gactccagtg cttgaaaaat gtattcggag aattgcaaga 480
 tcaaatcttg ataattgtatc gttggataag gaacttcctc cagaagtggc tgttcagata 540
 aaagagattc gccaaaaatc tcagccaaat gaggggtgaca ccgtcatttc agaccctgta 600
 catgagaaaa gggtcagaag aatccacagg gcaactggatt ctgatgatgt tgagcttgtg 660
 aagttgcttc ttaacgaatc tgagatcacc ttggatgatg ccaatgcatt gcactatgct 720
 gctgcttact gtgattcgaa agttgtttcg gagttgttag acttgagact tgccaacttg 780
 aatttgaaga attcgcgtgg atacacggca ctccatctgg ctgctatgag gagagagcca 840

```

gctattatca tgtgtctcct aaacaaagga gcagctgtat cacaattgac tgctgatggc 900
cagagtgcaa tgagtatctg ccggagggtta acaaggatga aagactacaa tacaaagatg 960
gagcaaggcc aagagtcaaa caaagacaga ttatgtattg atatattaga tagggagatg 1020
ataaggaaac ctatggcagt ggaagattct gtcacctcgc ctttggtggc tgacgatctt 1080
cacatgaagc ttctctacct tgaaaacaga gttgcatttg caagattatt ttttcctgca 1140
gaagcaaagg ttgcaatgca aattgcacaa gcagacacca caccagaatt tggcattggt 1200
cctgcagcta gcacttctgg aaaattgaag gaagtcgatc tgaacgagac accagtaaca 1260
caaaacaaaa ggctccgttc aagggtggat gcactcatga aaacagttga gctgggacgt 1320
cgctacttcc ctaactgctc gcagggtgctc gacaaatttc tggaggatga tttgcccgat 1380
agtcctgatg cactcgacct ccaaaatggc acttctgatg agcaaaatgt taaaaggatg 1440
cggttctgtg agttaaaagga ggatgtgcgc aaggcattca gcaaagacag agctgataat 1500
agcatgtttt ctatcttgtc atcttcatcg tcactcttcgc cactctccaa ggttgcaaaag 1560
aaatgacaga agttttgtaa caaatttccg ctctctgatg tactgggaca agagatatcg 1620
atcaatagac ctgtatagtc ttacagtggg ataacaatta gatatcgaag ctctctcgaa 1680
tattagaagg tgctgttctg ggctgcactc agctggttta tgggacccat gcggtgaaac 1740
tggcaaaaga aaaccagctg attagaggct ccaaagcagt gtctctcgtg aatatgtttg 1800
tagcattctg ttttgttcag gatggctata atgataaaat cttttcaata gatatatagc 1860
taattgtctc gtaaaaaaaa awaaaaaaa aaaagggggg gcccggtacc caattcgccc 1920
tatagtgaat cgtattacgc gcgctcactg gccgtcggtt tacaacgtcg tgactgggaa 1980
aaccttggtg ttacccaact taatcgctt gcagcacatc cccctttcgc cagctggcgt 2040
aatagcgaag aggccgcac cgatcgccc 2069

```

<210> 6
 <211> 455
 <212> PRT
 <213> *Oryza sativa*

<400> 6
 Asp Ala Phe Leu Ser Leu Leu Gly Tyr Leu Tyr Thr Gly Lys Leu Arg
 1 5 10 15
 Pro Ala Pro Asp Asp Val Val Ser Cys Ala Asp Pro Met Cys Pro His
 20 25 30
 Asp Ser Cys Pro Pro Ala Ile Arg Phe Asn Val Glu Gln Met Tyr Ala
 35 40 45
 Ala Trp Ala Phe Lys Ile Thr Glu Leu Ile Ser Leu Phe Gln Arg Arg
 50 55 60
 Leu Leu Asn Phe Val Asp Lys Thr Leu Val Glu Asp Val Leu Pro Ile
 65 70 75 80
 Leu Gln Val Ala Phe His Ser Glu Leu Thr Pro Val Leu Glu Lys Cys
 85 90 95
 Ile Arg Arg Ile Ala Arg Ser Asn Leu Asp Asn Val Ser Leu Asp Lys
 100 105 110
 Glu Leu Pro Pro Glu Val Ala Val Gln Ile Lys Glu Ile Arg Gln Lys
 115 120 125
 Ser Gln Pro Asn Glu Gly Asp Thr Val Ile Ser Asp Pro Val His Glu
 130 135 140
 Lys Arg Val Arg Arg Ile His Arg Ala Leu Asp Ser Asp Asp Val Glu
 145 150 155 160
 Leu Val Lys Leu Leu Leu Asn Glu Ser Glu Ile Thr Leu Asp Asp Ala
 165 170 175

Asn Ala Leu His Tyr Ala Ala Ala Tyr Cys Asp Ser Lys Val Val Ser
 180 185 190
 Glu Leu Leu Asp Leu Arg Leu Ala Asn Leu Asn Leu Lys Asn Ser Arg
 195 200 205
 Gly Tyr Thr Ala Leu His Leu Ala Ala Met Arg Arg Glu Pro Ala Ile
 210 215 220
 Ile Met Cys Leu Leu Asn Lys Gly Ala Ala Val Ser Gln Leu Thr Ala
 225 230 235 240
 Asp Gly Gln Ser Ala Met Ser Ile Cys Arg Arg Leu Thr Arg Met Lys
 245 250 255
 Asp Tyr Asn Thr Lys Met Glu Gln Gly Gln Glu Ser Asn Lys Asp Arg
 260 265 270
 Leu Cys Ile Asp Ile Leu Asp Arg Glu Met Ile Arg Lys Pro Met Ala
 275 280 285
 Val Glu Asp Ser Val Thr Ser Pro Leu Leu Ala Asp Asp Leu His Met
 290 295 300
 Lys Leu Leu Tyr Leu Glu Asn Arg Val Ala Phe Ala Arg Leu Phe Phe
 305 310 315 320
 Pro Ala Glu Ala Lys Val Ala Met Gln Ile Ala Gln Ala Asp Thr Thr
 325 330 335
 Pro Glu Phe Gly Ile Val Pro Ala Ala Ser Thr Ser Gly Lys Leu Lys
 340 345 350
 Glu Val Asp Leu Asn Glu Thr Pro Val Thr Gln Asn Lys Arg Leu Arg
 355 360 365
 Ser Arg Val Asp Ala Leu Met Lys Thr Val Glu Leu Gly Arg Arg Tyr
 370 375 380
 Phe Pro Asn Cys Ser Gln Val Leu Asp Lys Phe Leu Glu Asp Asp Leu
 385 390 395 400
 Pro Asp Ser Pro Asp Ala Leu Asp Leu Gln Asn Gly Thr Ser Asp Glu
 405 410 415
 Gln Asn Val Lys Arg Met Arg Phe Cys Glu Leu Lys Glu Asp Val Arg
 420 425 430
 Lys Ala Phe Ser Lys Asp Arg Ala Asp Asn Ser Met Phe Ser Ile Leu
 435 440 445
 Ser Ser Ser Ser Ser Ser
 450 455

<210> 7
 <211> 1052
 <212> DNA
 <213> Triticum aestivum

<400> 7
 gcacgagcag ggccaagagt caaataaaga taggatgtgc attgacatcc tagagagggga 60
 gatgatgagg aatcctatga cagcggaaga ttctgtcacc tcacctttat tggctgatga 120
 tcttcacatg aaactaagct acctggaaaa cagagtcgag ttcgcaagac tgttcttccc 180
 tgctgaagcc aaggttgcca tgcaaatgac acaagcagac gtcacaccag aagttggtgg 240
 tttttctgca gcaagtactt ctggtaaact gagggaaagtc gatctgaatg agacgccagt 300
 aacaaaaaac aaaaggctgc gttcaagggt ggatgcacta gcgaaaaacag tggaaactggg 360
 ccgtcggtag ttcccaaact gctcgcaggt gctcgacaaa ttcttggaag atggcctgcc 420
 tgatggcctt gatgcgttcc agcagcaaaag cggcaccgct gatgagcaac aggtgaagaa 480
 gatgcgcttc tgcgaggtga aggaggaagt gcgcaaaagca tacagcaaaag acacggccga 540
 taacagcatg ttttcggccc tgcgtcctc ctctcgtcc tcggcgatga agtgaaggta 600
 ctgtaacagg ctgttttctc gagatgtcag ggctaaagag ggatcgtctg tcatgcgcat 660
 gtatagtgtc caccatcgtg taaaaccgaa tatgaacatg aaaggaggcc ccaaaatagt 720
 agaagatgat atatactttg ctggacttgg agtttgttgg agaaggctgt gccatcccat 780
 cccagattcc caatatcaat ttcccatgct ggttgcaag acggagcgtt ggatcatcca 840
 gtttcgacgc tatgcagtcg tgcagcctgc tgtgtttgtt tcgcatagct gcaatactta 900
 tatgtttaat aatactagag agtagtaggc aattgaggct gtagcggaag ttggaacctta 960
 ccttaatgta agtgaaaggg gacagttgcc ctttgcgaa ctgttgttat caatacatag 1020
 ttgattttcg taaaaaaaaa aaaaaaaaaa aa 1052

<210> 8
 <211> 193
 <212> PRT
 <213> Triticum aestivum

<400> 8
 Glu Gln Gly Gln Glu Ser Asn Lys Asp Arg Met Cys Ile Asp Ile Leu
 1 5 10 15
 Glu Arg Glu Met Met Arg Asn Pro Met Thr Ala Glu Asp Ser Val Thr
 20 25 30
 Ser Pro Leu Leu Ala Asp Asp Leu His Met Lys Leu Ser Tyr Leu Glu
 35 40 45
 Asn Arg Val Ala Phe Ala Arg Leu Phe Phe Pro Ala Glu Ala Lys Val
 50 55 60
 Ala Met Gln Ile Ala Gln Ala Asp Val Thr Pro Glu Val Gly Gly Phe
 65 70 75 80
 Ser Ala Ala Ser Thr Ser Gly Lys Leu Arg Glu Val Asp Leu Asn Glu
 85 90 95
 Thr Pro Val Thr Lys Asn Lys Arg Leu Arg Ser Arg Val Asp Ala Leu
 100 105 110
 Ala Lys Thr Val Glu Leu Gly Arg Arg Tyr Phe Pro Asn Cys Ser Gln
 115 120 125
 Val Leu Asp Lys Phe Leu Glu Asp Gly Leu Pro Asp Gly Leu Asp Ala
 130 135 140
 Phe Gln Gln Gln Ser Gly Thr Pro Asp Glu Gln Gln Val Lys Lys Met
 145 150 155 160
 Arg Phe Cys Glu Val Lys Glu Asp Val Arg Lys Ala Tyr Ser Lys Asp
 165 170 175

WO 00/28036

PCT/US99/25953

Thr Ala Asp Asn Ser Met Phe Ser Ala Leu Ser Ser Asn Ser Ser Ser
180 185 190

Ser